המילון ההיסטורי ללשון העברית

מיסודם של

האקדמיה ללשון העברית ומוסד ביאליק

דוגמאות לקונקורדנציה ולאוספי מלים

של

"ברייתא דמלאכת המשכן"

ניסיון של אוטומאציה בחקר העברית במכונות IBM

ירושלים תשכ"א

הקדמת המערכת — זאב בן-חיים

התקנת הטקסט — גד בן-עמי צרפתי

תכנון עבודת המכונות — מרטין צוריאנו

בירור האפשרויות להשתמש בשיטת האוטומאציה - היינו במכונות חשמליות ואלקטרוניות לעיבוד נתונים - לצורך כינוסו של החומר למילון ההיסטורי, ובייחוד לצורך יצירת כרטסת-המובאות, הוא מן העניינים העיקריים, שנתנה עליו מערכת המילון ההיסטורי את דעתה בשלב זה של פעולתה. *) שיטה זו בעבודה הבלשנית ללשונות הנכתבות באלף-בית הלאטיני כבר יצאה מכלל גישושים וניסיונות, והיא הולכת וכובשת לה מקום, בפרט בתחום המילונאות, ואילו לגבי הלשון העברית אף לכלל התחלות עדיין לא נכנסה כניסה של ממש.

ההישגים שהושגו בשיטה חדשה זו לגבי הלשונות האמורות אין אתה יכול להעביר אל העברית מאליהם. הכתב העברי, שאינו כולל באלף-בית שלו סימנים מוגדרים לתנועות, וכן עצם השוני שבמבנה העברית לעומת הלשונות שנתנסו בשיטת האוטומאציה, מעלים שאלות מיוחדות לה ומחייבים למצוא פתרונות מכוונים לה. כדי לקדם את ליבונן של השאלות הכרוכות בהכנת החומר למלון ההיסטורי בשיטה זו וכדי לסלול את הדרך ליישובן נערך הניסיון הזה.

לצורך הניסיון נבחר אחד המקורות של המילון ההיסטורי, (טעמי הבחירה מפורשים להלן) והוא "ברייתא דמלאכת המשכן". ממקור זה מוצה <u>כל</u> החומר הלשוני ומוין בכמה כיוונים, עד שהמיון עשוי לשמש לא רק את המילונאי, אלא אף את חוקר לשונו וסגנונו של המקור. מטרתו של הניסיון כפולה: א) להעמיק חקר בבעיות האוטומאציה לגבי העברית לפרטי פרטיהן;

---

*) ע' "לשוננו" כרך כג עמ' 115, ושם כרך כה עמ' 81-95.

ב) לעמוד, במידת האפשר, על היקף ההוצאות, שהאוטומאציה בכינוסו ובמיונו של החומר הלשוני דורשת.

הניסיון נעשה בדרך "התקנה מוקדמת" (פְּרֵיאֶדִיציה) של המקור. אופן ההתקנה והבחנים שנקבעו לרישום שינויי גרסאות לצורך ניסיון זה מוסברים להלן. ההתקנה המוקדמת שננקטה אינה שלמה, שכן הוצא מכללה הסיווג הבלשני של כל מלה. התקנה מוקדמת מושלמת תובעת תחילה יישובן של שאלות-מבנה רבות, אבל לא היה טעם לאחר את הניסיון עד לאותה השעה. תקוותנו היא, שהניסיון כמות שהוא יועיל לקדם את יישובן של השאלות, שמערכת המילון מבקשת ליישבן קודם ליצירת כרטסת-המובאות.

הכרטסות, הכוללות את כל החומר שמוצא מן המקור המדובר, שמורות במערכת המילון ההיסטורי.*) חוברת זו באה אך להדגים את הניסיון ולפרסמו ברבים לטובת כל המעוניינים בדבר.

טובה כפולה ומכופלת אנו מחזיקים לחברת IBM בישראל על שנענתה לבקשת מערכת המילון ההיסטורי לסייע לה בביצוע הניסיון הזה. חברת IBM העמידה לרשות מערכת המילון את מומחיה ומכונותיה שלא על מנת לקבל פרס.

בביצוע הניסיון טרחו המהנדס מר יוסף כהן, מר מרטין צוריאנו ומר פיליפ עין-דור - מטעם IBM ; מן העובדים המדעיים של המילון: ד"ר גד בן-עמי צרפתי, בהתקנת הטקסט ובכל שאר תהליכי הביצוע, ומר ראובן מירקין סייע עמו.

החוברת מוגשת לקונגרס העולמי השלישי למדעי היהדות, העומד להתכנס בירושלים בימים י"ב - י"ט

---

*) הכתובת: האקדמיה ללשון העברית, ירושלים,
קרית האוניברסיטה, ת"ד 1033.

באב תשכ"א, והיא אף עשויה לשמש הדגמה להרצאתו של האב פרופ' רוברטו בוזה: "התפתחות תורת האינפורמציה בעולם והשפעתה על חקר הלשונות".

-- -- -- --

התקנת הטקסט

1) בחירת היחידה הספרותית וקביעת הנוסחה

א) לניסיון זה של עיבוד טקסט עברי על ידי אוטומאציה נבחרה ה"ברייתא דמלאכת המשכן", בעיקר משני טעמים אלו:

1. אורכה מספיק כדי להפיק מסקנות לגבי כמות העבודה שתידרש לעיבוד המקורות של המילון בשיטה זו;
2. הברייתא נמסרה לנו בכתבי יד באופן יחסי עתיקים ומדוייקים.

אף לקחנו בחשבון כגורם חיובי לבחירתנו, שהברייתא יצאה לאור במהדורה ביקורתית (על ידי מאיר איש-שלום, וינה תרס"ח/1908), ואפשר להסתייע בה בהתקנת הטקסט, אף כי אין לקבל את נוסחתה כבסיס לעבודה, בהיותה מלוקטת מתוך כתבי יד ודפוסים שונים.

הטקסט הותקן אפוא על פי:

1. קטע מהגניזה, מאמצע פרק ו' עד אמצע פרק י'; פורסם בגנזי שכטר, א, 377-383.
2. כ"י מינכן מס' 95 משנת ק"ג/1343; הוצאה פוטוגראפית ליידן 1912.
3. כ"י עפשטיין, נכתב בשאלוניקי בשנת רפ"ט/1529 (רמא"ש עמ' 3); אין מקום הימצאו היום ידוע; קטעים ממנו מובאים בפירושו של רמא"ש.
4. כ"י אוקספורד 370, לא קדם למאה י"ג (קטלוג נויבאור עמ' 25); העתק של כ"י זה נעשה על ידי מ' גרוסברג נמצא בבריטיש מוזיאום, וצילומו נמצא במכון לצילומי כתבי יד בירושלם.
5. כ"י פאריס מס' 353 מהמאה ט"ז; צילומו נמצא במכון

לצילומי כתבי יד בירושלם.

6. דפוס ויניציה משנת בשמ"ד/1602, בסוף הספר "משפטי שבועות" לר' האי גאון.

ממה שבדקנו נראה כי שש הנוסחאות האלה משקפות עריכה אחת, וההבדלים שביניהן אינם גדולים (פרט לכ"י פאריס).

ב) כיסוד הונח כ"י מינכן, השלם והעתיק מבין הנוסחאות שעמדו לרשותנו. כתב היד הועתק כמות שהוא, ורק במקרים המועטים שבהם טעות המעתיק גלויה, כגון שכתב משפט שאין בו מובן או שמובנו סותר את הידוע ואת המתקבל על הדעת, הוגה כתב היד על ידי הוספה, מחיקה, או שינוי - ורק אם הנוסחאות האחרות (כולן או מקצתן) מסייעות לנו. ההגהה צוינה בציון מיוחד, שלא לערבב אותה עם נוסחת כתב היד.

לא הגהנו את כתב היד במקום שחסר קטע שלם בפני עצמו (כגון: סוף פרק ו', ע' רמא"ש עמ' 41 שורות 4-6; ופרק ח', ע' גנזי שכטר א' עמ' 381 שורות 9-13), או אם אפשר ליישב - אפילו בדוחק - את הטקסט כמו שהוא, או אם הקיצור הוא פרי נטייה מכוונת של המעתיק, הכותב "משזר" במקום "שש משזר" (ב, 3) "טהרה" במקום "מקוה טהרה" (יב, 13), וכיוצא באלה.

במקום שיש בכתב היד חזרה על משפט שלם או על קטע שלם (כגון: פרק ו' שורות 7-13, ע' רמא"ש סוף עמ' 41 ראשית עמ' 42; פרק י"ג שורה 20) מחקנו את החזרה בלא לציין את הדבר.

ג) מלים וביטויים הבאים בכתב היד מקוצרים, וקיצורם מקובל, השארנו אותם כמות שהם (כגון: וגו׳ אע"פ שנ׳ ר׳), אבל השלמנו אותם במקרים אחרים; ההשלמה היא בדרך כלל בחזקת ודאי, פרט למקרים שבהם הכתיב של כתב היד אינו קבוע (כגון סיומת הרבים -ם או -ן). כמו כן השלמנו את המלים המעטות שלקו בכתב היד בחורים, בכתמים או במחיקות: השלמות אלו, שנעשו לפי הנוסחאות האחרות ולפי העניין, לא תמיד ודאיות הן. כל מלה שהושלמה סומנה בסימן מיוחד.

ד) ציינו בציון מיוחד את ההבאות מן המקרא, להבדיל בין לשון התנאים ללשון המקרא. במקום שההבאות הן פסוקים משובשים, או שאינן פסוקים ממש אלא צירופים של לשונות מן המקרא, לא תיקנו דבר. קיצורים בהבאות מן המקרא השלמנו, ולא צוין דבר.

ה) שאפנו ללוות את הטקסט בשינויי גרסאות מתוך כתבי היד האחרים, במידה שהשינויים הם בעלי ערך לשוני-מילוני, כדי לבדוק את אפשרות השילוב של שינויי הגרסאות בדרכי האוטומאציה. אחרי שערכנו רשימה שלמה של שינויים לגבי כשליש הברייתא (פרקים ו׳-י׳), קבענו על פיה בחנים לברירת שינויים אחדים לעיבוד האוטומאטי, ולפי אותם הבחנים ליקטנו שינויי גרסאות גם מיתר חלקי הברייתא. בחרנו בשינויים המשקפים שימוש שונה בבנייני הפועל (חפה - חיפה, ואולם ייתכן שאין זה אלא הבדל של כתיב), או שימוש במלים נרדפות (של - בת, בהוראת כמות; פרקים - פעמים), או שימוש בדרכי תחביר שונים (שמשמשים - ומשתמשים), וכן כיוצא בהם.

2) <u>התקנת הטקסט לעיבוד האוטומאטי</u>

הברייתא הועתקה מן המהדורה הפוטוגראפית של כתב היד ונכתבה במכונת הכתיבה שורה כנגד שורה. ספרור הפרקים ניתן כמו שהוא בשאר הנוסחאות, מ-א׳ עד י"ד, ולא כמו שהוא בכתב יד מינכן, מ-א׳ עד י"ג, שבו כנראה נשתבש על ידי כתיבת המספרים בסוף כל פרק ולא בראשו. ספרור השורות מתחיל ב-1 בכל פרק. כל הטקסט חולק שתי חלוקות נפרדות:

א) <u>החלוקה לקטעים.</u> כל קטע שהוא משפט אחד או ביטוי קצר המובן מתוך עצמו, ואורכו אינו עולה על תשע מלים; מספר המלים נקבע מטעמים טכניים. כל קטע נושא את מספר השורה שבה הוא מתחיל, בתוספת ספרה אם השורה מכילה את ההתחלות של קטעים אחדים (למשל: 16/0, 16/1, 16/2).

ב) <u>החלוקה לפסקות.</u> כל פסקה מכילה מעשרה עד חמישה עשר קטעים, והיא - עד כמה שאפשר - עניין שלם מבחינת התוכן. ספרור הפסקאות הוא רצוף בכל הברייתא, מ-1 עד 57.

אותיות השימוש מש"ה וכל"ב הפרדנו מן המלה בקו זקוף קטן, כדי שבכל אחת מהרשימות האלף-ביתיות יהיו המלים מסודרות לפי אותיותיהן הן ולא לפי אותיות השימוש הנלוות אליהן.

בטקסט נרשמו ספרות (כרגיל אחרי המלה, ולפעמים לפניה מסיבות טכניות) שהן ציונים מיוחדים. וזה פירושם:

1 - מלה מקוצרת שנעתקה בעינה.
2 - מלה מקוצרת שהושלמה בהעתקה.
3 - מלה לקויה או מטושטשת שהושלמה בהעתקה.
4 - מלה החסרה בטקסט ונוספה בהעתקה (על פי הנוסחאות האחרות).
5 - מלה מוטעית שתוקנה בהעתקה (כנ"ל) - ראה את גרסת כתב היד ברשימת שינויי הגרסאות.
6 - מלה מיותרת שיש למחקה (כנ"ל).
7 - מלה שיש לה שינוי, והוא רשום ברשימת שינויי הגרסאות.
8 - הבאה מן המקרא.
9 - מלה, שמספר אותיותיה עולה על שבע, ולא נכתבה בשלמות בכרטיסי-הטקסט, מטעמים טכניים.

3) <u>העיבוד האוטומאטי</u>

אחרי ההכנה הזאת נמסר הטקסט לעיבוד האוטומאטי, ובו עבר שלבים אלו:

א) <u>כרטיסי-טקסט (text-cards)</u>

הטקסט הועתק במכונה המנקבת, כל קטע על כרטיס לעצמו (או על כרטיס כפול אם הקטע עולה על שש מלים) בצירוף מראה המקום של הקטע בטקסט.

ב) <u>כרטיסי-מלים (word-cards)</u>

מכרטיסי-הטקסט נעשתה באופן אוטומאטי סדרה חדשה של כרטיסים מנוקבים, שכל אחד מהם נושא מלה אחת עם מראה מקומה (מראה מקום הקטע שאליו המלה שייכת בתוספת המספר הסידורי של המלה בקטע). על כל כרטיס נעתק (באותיות רגילות, ולא בניקוב) הקטע שאליו

המלה שייכת - וקיימת אף האפשרות (שלא נוצלה עתה) להעתיק עליו טקסט יותר ארוך, עד שלוש עשרה שורות. בשלב מאוחר יותר, נוקב בכל כרטיס-מלה המספר הסידורי של כרטיס-הערך המתאים. כרטיסי-המלים אף עשויים להצטרף לכרטסת-המובאות הכללית של המילון ההיסטורי. לסדרה זו נוספה רשימת שינויי הגרסאות: כל מלה שהיא שינוי גרסה נוקבה על כרטיס-מלה מיוחד, הנושא את מראה המקום של המלה המקבילה שבטקסט, בצירוף ציון המקור שממנו גרסה זו לקוחה. הציונים הם:

ג - קטע מן הגניזה, גנזי שכטר, א, 377-383.
מ - כתב יד מינכן 95.
ע - כתב יד עפשטיין, על פי המובאות של רמא"ש.
א - כתב יד אוקספורד 370, על פי ההעתקה של גרוסברג.
פ - כתב יד פאריס 353.
ד - דפוס ויניציה שס"ב.

ג) <u>כרטיסי-ערכים (main-cards)</u>

על פי כרטיסי-המלים נעשתה, בידי אדם, רשימת הערכים של הטקסט. עשיית רשימה זו פירושה: ריכוז של מלים שונות השייכות לערך אחד, מצד אחד, והפרדת המלים ההומוגראפיות, מצד שני. למשל: המלים "מביתו" "בתינו" "לבתיכם" נותנים את הערך "בית"; המלים "ישבנו" בשבתכם" "שב" - את הערך "ישב". ואילו צירוף האותיות "שתי", אם הוא מופיע בטקסט פעמים אחדות, עשוי להעמיד ארבעה ערכים שונים (ואף יותר): "שְׁתִי" - שם עצם, "שתים" (סמיכות), "שתה" (ציווי נקבה), "שית" (עבר מדבר).

על פי רשימת הערכים נוקבו כרטיסי-הערכים. בכרטיסים אלו נוקבו באופן אוטומאטי גם המספר הסידורי של הערך לפי סדר האלף-בית, ומספר ה"תדירות", היינו מספר הפעמים שהערך מופיע בטקסט לצורותיו השונות. בשלב זה נוקב מספר הערך בכרטיסי-המלים.

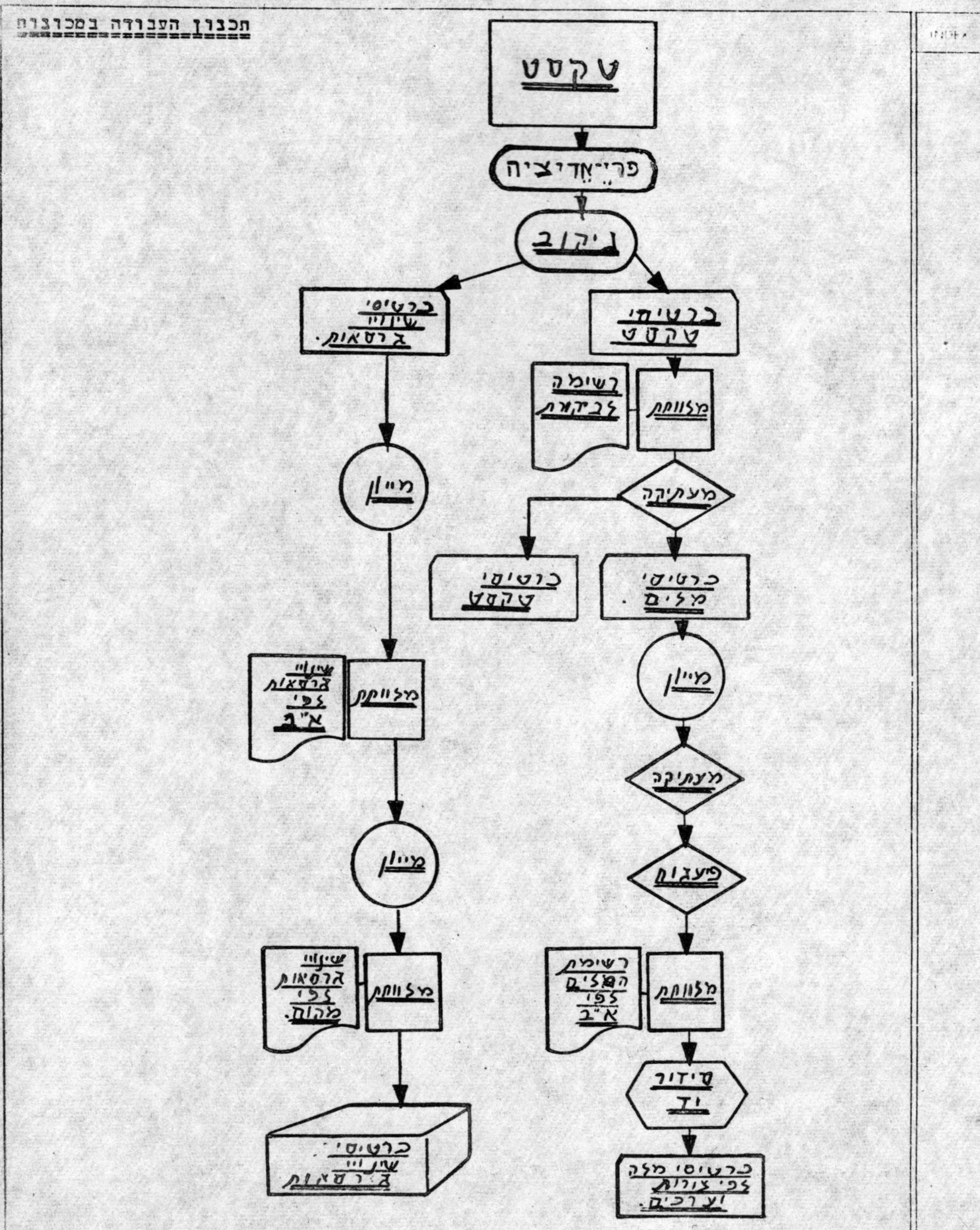
ד) כרטיסי-צורות (form-cards)

מתוך כרטיסי-המלים, מחולקים לפי הערכים, ייצרה המכונה את כרטיסי-הצורות: הם מכילים את כל הצורות השונות הנמצאות בחיבור - כל אחת רק פעם אחת. למשל: אם בחיבור נמצאות שש המלים "דבר" "דבר" "דברי" "דברים" "דברים" "דברים", הן נותנות שישה כרטיסי-מלים, ומהם מתקבלים שלושה כרטיסי-צורות: "דבר" "דברי" "דברים". בכל כרטיס-צורה מצטרפים לצורה: ציון תדירותה, ציון מקומה בסדר האלף-ביתי של כל הצורות שבטקסט, המספר הסידורי של הערך שאליו הצורה שייכת.

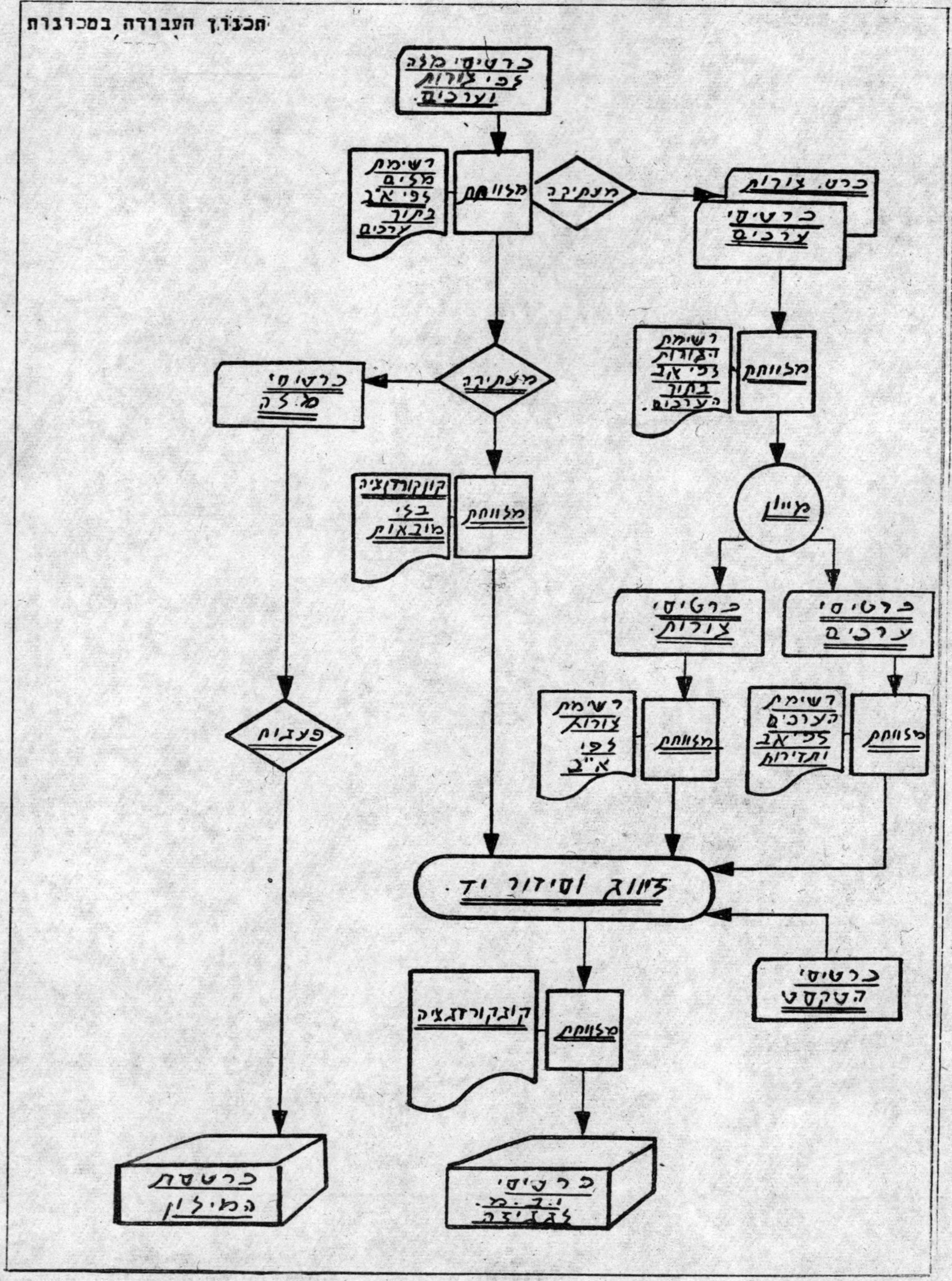
ה) רשימות מלים וקונקורדנציות

על ידי עיבוד אוטומאטי של ארבע סדרות הכרטיסים (כרטיסי-הטקסט, כרטיסי-המלים, כרטיסי-הערכים, כרטיסי-הצורות) נערכו רשימות שונות של מלים לפי בחנים שונים. דוגמאות של רשימות אלה ניתנות להלן.

FLOW CHART


תכנון העבודה במכונות
טקסט
פרי־אדיציה
ניקוב
כרטיסי שינויי גרסאות.
כרטיסי טקסט
רשימה לביקורת
מלווחת
מעתיקה
כרטיסי טקסט
כרטיסי מלים.
מיין
שינויי גרסאות לפי א"ב
מלווחת
מיין
מעתיקה
פיענוח
שינויי גרסאות לפי מהות.
מלווחת
רשימת המלים לפי א"ב
מלווחת
סידור יד
כרטיסי שינויי גרסאות
כרטיסי מלה לפי צורות וערכים.

FLOW CHART


תכנון העבודה במכונות
כרטיסי מלה לפי צורות וערכים.
מלווחת
רשימת מלים לפי א"ב בתוך ערכים
מעתיקה
כרט. צורות
כרטיסי ערכים
מלווחת
רשימת הצורות לפי א"ב בתוך הערכים.
מיון
כרטיסי צורות.
כרטיסי ערכים
מלווחת
רשימת צורות לפי א"ב
מלווחת
רשימת הערכים לפי א"ב ותדירות
מעתיקה
כרטיסי מלה
מלווחת
קונקורדנציה בלי מובאות
פענוח
זיווג וסידור יד.
כרטיסי הטקסט
מלווחת
קונקורדנציה
כרטסת המילון
כרטיסי י.ב.מ. לגניזה

ביאור הדוגמאות

1) **העתק הטקסט**

הטקסט מועתק במכונת כתיבה ומוכן לעיבוד האוטומאטי.

2) **רשימת כרטיסי-הטקסט**

כרטיסי-הטקסט, נעתקים בזה אחר זה, נותנים שוב את הטקסט בשלמותו. על יד כל שורה בא מראה-מקומה.
פירוש הספרות - משמאל לימין:
510 - ציון מספרי שנקבע ל"ברייתא דמלאכת המשכן"
2 ספרות - מספר הפסקה
2 ספרות - מספר הפרק
2 ספרות - מספר השורה
ספרה אחת - חלק של השורה

3) **רשימת המלים**

כל המלים של הטקסט לפי סדר האלף-בית, עם אותיות השימוש ועם ספרות-ציון.
טור 1 - המספר הסידורי של המלה
טור 2 - המלה
טור 3 - מראה מקום המלה (ע' רשימה 1)

4) **רשימת שינויי-הגרסאות**

שינויי-הגרסאות שנבחרו, בסדר האלף-בית.
טור 1 - מראה המקום: סדר הספרות כברשימה 1, בתוספת ספרה אחת המציינת את מקום המלה בשורה
טור 2 - שינוי הגרסה
טור 3 - ציון המקור שממנו נלקח שינוי הגרסה

5) **רשימת הצורות**

כל הצורות המופיעות בטקסט, לפי סדר האלף-בית של הערכים.
טור 1 - המספר הסידורי של הערך שאליו הצורה שייכת
טור 2 - הצורה
טור 3 - תדירות הצורה
טור 4 - תדירות הערך

6) **רשימת הערכים לפי סדר האלף-בית**

טור 1 - המספר הסידורי
טור 2 - הערך
טור 3 - התדירות

7) **רשימת הערכים לפי סדר תדירותם**

טור 1 - המספר הסידורי של הערך לפי סדר האלף-בית
טור 2 - הערך
טור 3 - התדירות
טור 4 - מספר הערכים בעלי תדירות שווה
טור 5 - המספר הסידורי של הערך לפי סדר התדירות

8) **רשימת הערכים והצורות**

הערכים סדורים לפי האלף-בית; אחרי כל ערך באות הצורות השייכות לו, אף הן לפי סדר האלף-בית.
טור 1 - המספר הסידורי של הערך בסדר האלף-בית
טור 2 - המספר הסידורי של הצורה בסדר האלף-בית של הצורות
טור 3 - הערך
טור 4 - הצורה
טור 5 - תדירות הצורה
טור 6 - תדירות הערך

9) **קונקורדנציה בלי מובאות**

טור 1 - המספר הסידורי של הערך בסדר האלף-בית
טור 2 - המספר הסידורי של המלה בסדר האלף-בית של המלים
טור 3 - הערך
טור 4 - המלה
טור 5 - מראה מקום המלה (ע' רשימה 3)
טור 6 - תדירות הערך

10) **קונקורדנציה**

כל הערכים בסדר האלף-בית, לפני כל ערך מספרו הסידורי ואחריו תדירותו. אחרי הערך באות הצורות השייכות לו (בסדר האלף-בית), על יד כל אחת רשומה תדירותה. אחרי כל צורה באות מובאותיה, עם מראי-מקומותיהן; פירוש הספרות, משמאל לימין: הפרק, השורה, החלק של השורה.

-- -- --

1

פרק 1

רבי
אומר עשר תרומות הן תרומה
גדולה ותרומת מעשר
וחלה ובכורים ותרומת
איל בזיר ותרומת לחמי
תודה ותרומת הימין ותרומת מדין ותרומת שקלים ותרומת
המשכן תרומה ותרום' מעשר וחלה ובכורים ותרום' איל בזיר ותרומת
לחמי תודה לכהנים ותרומת הארץ לכהנים וללוים ולנתינים ולמקדש
ולירושלם תרומת מדין לאלעזר הכהן תרומת שקלים לאדנים תרום'
המשכן שממנה היו עושין גופו ארכו שלשים אמה ורחבו עשר אמו'
וגבהו כרחבו ר' יוסי או' ארכו שלשים ואחת אמה כיצד מעמיד את
הקרשים היה בונה עשרים קרש בצפון ועשרים בדרום ושמנה במער'
אבל במזרח לא היה שם קרש אלא ארבעה עמודי שטים שעליהם
בותנין את הפרכת שנ' ונתת על ארבעה עמודי שטים והיה
עושה את האדנים חלולין וחורץ את הקרש מלמטה רביע מכאן
ורביע מכאן וחורץ חציו באמצע ועושה לו שתי ידות כמין שני
חסוקין המכניסן לתוך שני אדנים תחת הקרש האחד שני
הסיגין יוצאין מתוך הקרשים שנים לכל אחד ואחד שמשקיע
והכר בתוך הנקבה שנ' משלבת אשה אל אחותה דברי ר' נחמיה
ר' נחמיה אומ' אין ת"ל משולבות ומה ת"ל משולבות שעשה להם
שלבין כסולם המצרי היה חורץ את הקרש מלמעלן אצבע
מכאן ואצבע מכאן ונותנן בתוך טבעות של זהב כדי שלא יהו נפרדי'
זה מזה שנ' ויהיו תואמים מלמטה ויחדו יהיו תמים על ראשו את

2

5100101010	רבי4	אומר	עשר	תרומות	הז				
5100101011	תרומה	גדולה	ותרומת	מעשר	וחלה	ובכורים.			
5100101030	ותרומת	איק	נזיר	ותרומת	קחמי	תודה			
5100101050	ותרומת	היתיז	ותרומת	מדיז					
5100101051	ותרומת	שקקים	ותרומת	המשכז					
5100101060	תרומה	ותרומת2	מעשר	וחלה	ובכורים				
5100101061	ותרומת2	איק	נזיר	ותרומת	קחמי	תודה	קכהנים		
5100101070	ותרומת	הארץ	קכהנים	וקקוים	וקנתינים	וקמקדש	וקירושקם		
5100101080	תרומת	מדיז	קאקעזר	הכהז					
5100101081	תרומת	שקקים	קאדנים	תרומת2	המשכז	שמתנה	היו	עושיז	גופו
5100201090	ארכו	שקשים	אמה	ורחבו	עשר	אמות2	וגבהו	כרחבו	
5100201100	ר1	יוסי	אומר2						
5100201101	ארכו	שקשים	ואחת	אמה					
5100201102	כיצד	מעמיד	את	הקרשים					
5100201110	היה	צוחז	עשרים	קרש	בצפוז	ועשרים	בדרום	ושמנה	במערב2
5100201120	אבק	במזרח	קא	היה	שם	קרש			
5100201121	אקא	ארבעה	עמודי	שטים					
5100201122	שעקיהם	נוחניז	את	הפרכת					
5100201130	שנ1	ונתתה8	עק8	ארבעה8	עמודי8	שטים8			
5100201131	והיה	עושה	את	האדנים	חקוקיז				
5100201140	וחורץ	את	הקרש	מקמטה	רביע	מכאז	ורביע	מכאז	
5100201150	וחורצד	חציו	באמצע						
5100201151	ועושה	קו	שתי	ידות	כמיז	שני	קבחמוקיז		
5100201160	ומכניסז	קתוך	שני4	אדנים4					
5100201161	שנ14	שני8	אדנים8	תחת8	הקרש8	האחד8			
5100301162	ושני	הסיניז	יוצאיז	מתוך	הקרשים	שנים	קכק	אחד	ואחד
5100301170	שמשקיע	הזכר	בתוך	הנקבה					
5100301180	שנ1	משקבות8	אשה8	אק8	אחותה8				
5100301181	דברי	ר1	נחמיה	שר1	נחמיה	אומר2			

3

5104912271	הבקרים8	0739
5102105290	בר1	0740
5100301231	הבריח	0741
5100501370	והבריח8	0742
5100601441	שהבריחים	0743
5100601430	שהבריחים	0744
5100501330	והבריחים	0745
5100601401	והבריחים	0746
5100501321	הבריחים	0747
5105113081	הבריחים2	0748
5100601410	לבריחים8	0749
5100601420	לבריחים8	0750
5100601420	לבריחים8	0751
5102506301	הברית	0752
5102506290	הברית	0753
5102406250	ברית8	0754
5102506302	הברית8	0755
5102206040	בת	0756
5104712100	הבת	0757
5104812211	בת8	0758
5104712101	והבת8	0759
5104712071	בת8	0760
5104712110	בתים	0761
5100601410	בתים8	0762
5100601420	בתים8	0763
5104712110	בתים8	0764
5100601420	בתים8	0765
		112
5104211020	גבהו	0766
5104311060	וגבהו	0767

ת	האהל	510	05	12	0	4
ת	אורכה	510	04	02	1	2
ד	מלאחוריהם	510	14	06	0	4
ד	אחת	510	06	18	0	5
ד	ואחת	510	06	18	0	6
ת	אמנת	510	07	11	0	1
ת	ארבעה	510	13	15	1	4
ד	ארכן	510	06	18	0	2
ת	הבית	510	07	22	0	6
א	ברקים	510	11	15	0	3
ע	הברתיים	510	10	18	0	6
אפד	בת	510	06	16	1	2
ע	דולקין	510	10	32	0	2
א	דליקתן	510	10	32	0	2
א	9הבירותיים	510	10	18	0	6
ד	הוילאות	510	03	05	0	4
ע	הוצא	510	06	05	0	1
ע	הוצא	510	06	19	1	1
ע	הוצא	510	06	21	0	1
פא	והחריט	510	01	15	0	1
ת	היה	510	06	28	0	2
א	שהיה	510	10	32	0	1
ת	שהיו	510	06	27	1	2
ת	היו	510	10	29	0	7
אעד	הכרעה	510	10	07	1	6
אד	הכרעה	510	10	12	1	7
ת	המנורה	510	04	21	0	2
א	וילאות	510	02	08	1	4
עא	וילאות	510	03	05	0	4
ת	ולמשה	510	10	03	0	3
גד	ושני	510	07	13	0	1
ת	זהב	510	02	10	1	5
א	זהב	510	11	03	0	4
ע	חווקים	510	01	15	1	7
גאפ	וחיפה	510	07	04	1	1
געא	שחיפה	510	07	06	0	2
אד	החתה	510	14	12	0	5
ת	וחפוי	510	07	22	0	4
גאפ	יהו	510	07	08	1	6
ד	כאן	510	12	15	0	9
ע	כולי	510	11	17	0	2
גע	כופל	510	08	06	0	1
ג	כופל	510	08	09	1	1
אד	שכופלה	510	03	12	0	5
גאפד	כיצד	510	07	01	0	1
ת	ככר	510	01	32	0	2
ת	וככר	510	01	32	0	3
געתא	מככר	510	09	16	1	2
געאד	מככר	510	09	16	1	5
ת	שכל	510	01	34	0	2
ד	כלה	510	11	17	0	2

5

066	בהן	2	
066	בו	7	15
067	בבל	2	2
068	בדיו	1	
068	בדים	6	7
069	הבדיל	1	1
070	בהמתם	1	1
071	בא	2	
071	באה	9	
071	באות	1	
071	יבא	1	
071	יבאו	1	
071	תביא	4	18
072	בין	2	
072	ביניהם	2	4
073	בית	14	
073	בתים	3	17
074	בכורים	2	2
075	בוקטין	1	1
076	בן	6	
076	בני	8	
076	בניו	2	16
077	בניתן	1	1
078	בנין	1	1
079	בערה	1	1
080	בצלאל	2	2
081	בקר	3	
081	בקרים	2	5
082	בר	1	1

55	ארון	25
56	ארר	5
57	ארר	41
58	ארץ	2
59	ארר	3
60	אשה	3
61	אשם	1
62	אשר	1
63	אשר	2
64	את	81
65	אתה	7
66	ב	15
67	בבל	2
68	בד	7
69	בדל	1
70	בהמה	1
71	בוא	18
72	בין	4
73	בית	17
74	בכורים	2
75	בלט	1
76	בן	16
77	בנימין	1
78	בנין	1
79	בער	1
80	בצלאל	2
81	בקר	5
82	בר	1
83	בריח	11
84	ברית	4
85	בת	6

7

		1	
23	27	קרש	477
		1	
24	26	ארבע	50
25	26	חמשים	169
26	26	משכן	342
		3	
27	25	ארון	55
		1	
28	24	יצא	208
29	24	כאן	224
30	24	שלחן	523
		3	
31	23	ל	265
		1	
32	22	מה	291
33	22	מנורה	320
34	22	שלוש	522
		3	
35	21	אלא	28
36	21	זה	143
37	21	מזבח	296
38	21	שתים	552
39	21	תל	572
		5	
40	20	עמוד	410
		1	
41	19	אהל	7
42	19	אחת	19
43	19	משה	339
		3	
44	18	בוא	71
		1	
45	17	בית	73
46	17	יריעה	213
47	17	פרח	437
48	17	פרכת	438
49	17	קלע	465
50	17	תרומה	582
		6	
51	16	בן	76
52	16	מעלה	325
53	16	צפון	451
		3	
54	15	ב	66
55	15	גביע	90
56	15	דרום	119
57	15	מאה	286

8

066	ב				15
		100	בה	3	
		101	בהם	3	
		103	בהן	2	
		104	בו	7	
067	בבל				2
		097	בבל	2	
068	בד				7
		098	בדיו	1	
		099	בדים	6	
069	בדל				1
		175	הבדיל	1	
070	בהמה				1
		102	בהמתם	1	
071	בוא				18
		094	בא	2	
		095	באה	9	
		096	באות	1	
		269	יבא	1	
		270	יבאו	1	
		413	מביא	4	
072	בין				4
		106	בין	2	
		107	ביניהם	2	
073	בית				17
		108	בית	14	
		125	בתים	3	
074	בכורים				2
		109	בכורים	2	
075	בלט				1
		105	בולטין	1	
076	בן				16
		110	בן	6	
		111	בני	8	
		112	בניו	2	
077	בנימין				1
		113	בנימן	1	
078	בנין				1
		114	בנין	1	
079	בער				1
		115	בערה	1	
080	בצלאל				2
		116	בצלאל	2	
081	בקר				5

9

Count	Reference	Form	Root	Serial	Root no.
	510 06 01 42 0 3	בתים8		0763	
	510 06 01 42 0 7	בתים8		0765	
2			בכורים		074
	510 01 01 01 1 6	ובכורים		0711	
	510 01 01 06 0 5	ובכורים		0712	
1			בלט		075
	510 28 07 18 0 4	בולטין		0692	
16			בן		076
	510 38 10 07 0 2	בן		0714	
	510 25 06 26 1 3	בן		0715	
	510 36 09 16 0 3	בן		0716	
	510 56 14 09 2 3	בן		0717	
	510 29 07 23 1 3	בן		0718	
	510 57 14 17 1 3	בן		0719	
	510 50 13 01 1 1	בני		0720	
	510 50 13 04 0 1	בני		0721	
	510 51 13 07 0 1	בני		0722	
	510 53 13 21 1 2	בני		0723	
	510 45 11 22 0 2	בני		0724	
	510 50 13 02 0 3	בני8		0725	
	510 53 13 23 0 5	בני8		0726	
	510 53 13 21 0 5	בני8		0727	
	510 51 13 10 0 4	ובני		0728	
	510 53 13 22 0 7	ובני8		0729	
1			בנימין		077
	510 50 13 06 1 7	ובנימין		0730	
1			בנין		078
	510 30 07 29 0 2	בבנין		0731	
1			בער		079
	510 40 10 25 1 4	לבערה8		0732	
2			בצלאל		080
	510 26 07 01 0 3	בצלאל		0733	
	510 37 10 01 0 3	בצלאל		0734	
5			בקר		081
	510 49 12 27 1 2	הבקר8		0735	
	510 49 12 25 0 3	בקר8		0736	
	510 49 12 23 0 6	בקר8		0737	
	510 49 12 25 0 5	לבקרים8		0738	
	510 49 12 27 1 5	הבקרים8		0739	
1			בר		082
	510 21 05 29 0 5	ברן		0740	
11			בריח		083

10

7 27 0	בית קדשיך הקדשים2 שעשה קו6 שלמה
4 09 1	ובית עולמים ארכו שעים אמה
7 23 1	ר1 יהודה2 בן לקיש אומר2 במקומו בבית קדשיך הקדשים
4 08 1	כר היו מונחין בבית עולמים2
7 12 0	חולק את הבית עשר אמות על שער אמות
7 11 0	ארונ5 היה נתונ5 בתור הבית
4 10 0	הא למדת שאהל רביעי7 של בית עולמים
4 17 1	ומזבח הזהב היה נתון בתור הבית
7 15 1	לפיכך לא ננעלו דלתות בית קדשיך הקדשים מעולם
12 24 0	לעזרה פונה לימין להר הבית פונה
4 18 0	וחולק את הבית ומחציו לפנים מכוון כנגד הפרכת
5 29 1	קרפף יותר מבית סאתים כחצר המשכן
14 13 1	שנ14 לכל8 עיני8 כל8 בית8 ישראל8
13 07 1	שנ1 ונשיא8 בית8 אב8 למשפחות8 מררי8

בתים 3

1 41 0	ואת8 הקרשים8 תצפה8 זהב8 בתים8 לבריחים8
1 42 0	שאין תלו בתים8 לבריחים8 ומה תלו בתים8 לבריחים8
1 42 0	שאין תלו בתים8 לבריחים8 ומה תלו בתים8 לבריחים8

74 בכורים 2

בכורים 2

1 01 1	תרומה גדולה ותרומת מעשר וחלה ובכורים
1 06 0	תרומה ותרומת2 מעשר וחלה ובכורים

75 בלט 1

בולטין 1

7 18 0	הא4 כיצד4 היו בולטין לפרכת7

76 בן 16

בן 6

10 07 0	איסי בן יהודה2 אומר2
6 26 1	ר1 יהודה2 בן לקיש אומר
9 16 0	ר1 יהושע בן קרחה אומר2
14 09 2	ר1 שמעון בן יוחאי אומר2
7 23 1	ר1 יהודה2 בן לקיש אומר2 במקומו בבית קדשיך הקדשים
14 17 1	ר1 שמעון בן עזאי אומר2 ממזבח2 הקטרת2

בני 8

13 01 1	בני קהת היו עומדין בדרום
13 04 0	בני גרשון היו משמרים2 במערב2
13 07 0	בני מררי היו משמרין בצפונ5
13 21 1	נכנסו בני קהת ופרקו את הקדש
11 22 0	ששו בני האולה גבוה שער אמות
13 02 0	שנ1 משפחות8 בני8 קהת8
13 23 0	שנ1 ונסע8 דגל8 מחנה8 בני8 דן8
13 21 0	שנ1 ונסע8 דגל8 מחנה8 בני8 אפרים8

בניו 2
13 10 0 לתזרח משה ואהרז ובניו וכלה8 אהרז8 ובניו8
13 22 0 וטענו אותו בכתף ענו 1

77 בנימיז 1

בנימז 1
13 06 1 חוצה2 להם עלעה2 עבטים אפרים2 ומנשה2 ובנימז

78 בניז 1

בניז 1
7 29 0 אבל בבניז האחרוז ילא היה5 עם כותל

79 בער 1

בערה 1
10 25 1 מנורת8 זהב8 פונרותיהם לבערה8 כמשפט8 לפני8 הדביר8 זהב8 סגור8

80 בצלאל 2

בצלאל 2
7 01 0 לבאיזהצד עשה2 בצלאל את הארוז
10 01 0 כיצד עשה בצלאל את המנורה2

81 בקר 5

בקר 3
12 27 1 ואת8 הבקר8 ושני8 טורי8 הבקרי8 מוצקים8 במצקתו8
12 25 0 ענו 1 ודמות8 בקר8 מתחת8 לבקרים8 סביב8
12 23 0 ענו 41 עומד8 על8 ענים8 עשר8 בקר8

בקרים 2
12 25 0 ענו 1 ודמות8 בקר8 מתחת8 לבקרים8 סביב8
12 27 1 ואת8 הבקר8 ושני8 טורי8 הבקרים8 מוצקים8 במצקתו8

82 בר 1

בר 1
5 29 0 תנאו אמר2 רו יוסי ברו יהודה2

ברית 2

חלוקת הערכים לפי התדירות

תדירות	מספר הערכים	תדירות	מספר הערכים	תדירות	מספר הערכים
120	1	35	1	15	7
110	1	31	1	14	8
102	1	29	1	13	3
99	1	27	1	12	7
81	1	26	3	11	14
69	1	25	1	10	8
61	1	24	3	9	8
60	1	23	1	8	12
56	1	22	3	7	20
52	1	21	5	6	22
45	2	20	1	5	20
44	1	19	3	4	38
43	1	18	1	3	51
41	2	17	6	2	78
37	2	16	3	1	239
36	1				

סיכומים מספריים

מספר המלים של הטקסט	3729
מספר הצורות	899
מספר הערכים	588

חלוקת המלים לפי האלף-בית

א	654	ל	109
ב	110	מ	468
ג	48	נ	140
ד	67	ס	28
ה	197	ע	296
ו	9	פ	102
ז	75	צ	46
ח	163	ק	100
ט	65	ר	114
י	162	ש	450
כ	201	ת	125

חלוקת הערכים לפי האלף-בית

א	66	ל	21
ב	21	מ	64
ג	14	נ	23
ד	23	ס	16
ה	18	ע	36
ו	1	פ	23
ז	11	צ	7
ח	26	ק	26
ט	9	ר	19
י	35	ש	60
כ	42	ת	25